Aggregated Community Question Answering

Alton Y. K. Chua and Snehasish Banerjee

***Abstract*—The objective of this paper is two-fold objective. First, it introduces ACQUA, a conceptual prototype for an Aggregated Community QUestion Answering site, which provides question clustering and answer prediction functionalities across an aggregation of multiple CQAs. Second, using a combination of questionnaires and group interviews, it conducts an evaluation of users' behavioral intention to adopt ACQUA based on four parameters, namely, perceived usefulness, perceived ease of use, competence, and usability. Results indicate that the behavioral intention to adopt ACQUA was generally promising. In particular, ACQUA received favorable responses from most participants in terms of perceived usefulness, perceived ease of use, and competence. In terms of usability however, the consensus was not completely in favor of ACQUA. Three implications arising from the findings are discussed. Finally, the paper concludes with notes on limitations and future work.**

***Index Terms*—community question answering, behavioral intention to adopt, question clustering, answer prediction**

I. Introduction

WITH the advent of Web 2.0, social media comprise one of the fastest growing segments on the Internet. A popular form of social media application that has gained widespread popularity of late as an information seeking channel includes community question answering sites (CQAs), which '*leverage user interactions and user-generated content in order to satisfy the information needs of the community members*' [1, p 116]. They represent dedicated avenues for users to exchange information by asking questions to the online community, and answering questions posted by others in the community.

Researchers from various disciplines such as information science, information retrieval, human computer interaction, and computer science are trying to grasp the depth and breadth of CQAs [2]. Two prominent themes of CQA research include question clustering and answer prediction. Clustering has often been done using analysis of hyperlinks and cross-references [3]. Two commonly used approaches involve bipartite [4], and tripartite graph-based clustering [5]. On the other hand, three approaches often followed for answer prediction deal with the use of social features [6], textual features [7], and content-appraisal features [8].

As a part of the larger ongoing project, this paper attempts to extend and assimilate these two disparate strands of CQA research by introducing ACQUA, a conceptual prototype for an Aggregated Community QUestion Answering site. Aggregating the corpora of four CQAs, namely, Answerbag, Askville, WikiAnswers and Yahoo! Answers, ACQUA represents a system that incorporates both question clustering and answer prediction functionalities. Question clustering is accomplished using a quadripartite graph-based clustering approach [9]. Answer prediction is done using a combination of social, textual, and content-appraisal features [10]. Specifically, the objective of this paper is two-fold. First, it introduces ACQUA, an aggregated CQA site. Second, using a combination of questionnaires and group interviews, it evaluates users' behavioral intention to adopt ACQUA.

The remainder of the paper is structured as follows. The next section presents a brief literature on question clustering, answer prediction, and behavioral intention to adopt. This is followed by the design overview of ACQUA. Next, the procedures for evaluating behavioral intention to adopt ACQUA are explained. Later, the paper highlights the results, and offers a discussion. Finally, it concludes with notes on limitations and future work.

II. Literature Review

A. Question Clustering and Answer Prediction

Clustering of similar online content has often been done using analysis of hyperlinks and cross-references [3], whereby both interacting agents (eg. users) and subjects of interaction (eg. questions) are represented using distinct types of nodes. Bipartite graph-based clustering approach that deals with two types of nodes has been used for applications such as ontology mapping [4], and query clustering [3]. Tripartite graph-based clustering approach, associated with three types of nodes, has been used for applications such as collaborative tagging [11], and image clustering [5]. Driven by these, ACQUA uses a quadripartite graph-based approach for question clustering. The quadripartite graph consists of four types of nodes, namely, questions, answers, askers, and answerers. This approach has been shown to out-perform baseline question clustering algorithms such as bipartite graph-based approach that uses only two types of nodes, that of questions and answers [9].

Prediction of high quality answers in CQAs has often been done using social features that are associated with the community aspects of users. Previous studies on answer quality and best answer prediction have consistently stressed on social features such as answerers' authority, and

Manuscript received December 5, 2012; revised January 21, 2013. This work is supported by Nanyang Technological University Research Grant AcRF Tier 1 (RG58/09).

Alton Y. K. Chua is with Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore 637718 (phone: +65-6790-5810; e-mail: altonchua@ntu.edu.sg).

Snehasish. Banerjee is with Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore 637718 (e-mail: snehasishb@ntu.edu.sg).

users' endorsement or disapproval for answers [6], [7]. A second approach has used textual features as proxies for answer quality in CQAs. Some commonly used textual features include answer length, number of unique words, and non-stop word overlap in answers [7]. A third approach lies in the use of content appraisal features to predict answer quality. Content-appraisal features widely cited by scholars include accuracy, completeness, and reasonableness [8], [12]. For the purpose of ACQUA, a combination of social features, textual features, and content-appraisal features are used for prediction of best answers in response to a given question [10].

B. Behavioral Intention to Adopt

Behavioral intention to adopt can be defined as the inclination of users to embrace a new technology, service, application or system [13]. Being highly influenced by the cognitive psychology of users, behavioral intention to adopt is increasingly deemed an essential condition before the effective implementation of any information technology system or application [14]. Several models have been devised to examine behavioral intention to. Some of these include the Theory of Reasoned Action (TRA), the Theory of Planned Behavior (TPB), and the Technology Acceptance Model (TAM).

Drawing collaboratively from such models, this paper evaluates behavioral intention to adopt ACQUA based on four parameters. These are perceived usefulness, perceived ease of use, competence, and usability [15], [16], [17]. Perceived usefulness refers to the extent to which users believe in the effectiveness of an application in fulfilling its intended purpose [15]. Perceived ease of use denotes the extent to which users conceive an application as user-friendly with a smooth learning curve, and that it can be used with minimal effort [17]. Competence refers to achieving a match between the skills of users and the challenges in task accomplishment, so that users' experience would neither be too monotonous nor too stressful [18]. Usability refers to the visibility, flexibility, ease of use of the system, and is regarded as a measure of its ability to improve users' task performance in such a way that can result in the system being used more frequently [16], [19].

III. ACQUA – Design Overview

The design of ACQUA aims to provide users with easy access to clustered similar questions, and predicted best answers by aggregating the corpora of four CQAs, namely, Answerbag, Askville, WikiAnswers, and Yahoo! Answers. ACQUA is intended to serve the purpose of an easy-to-use information seeking site for all Internet users, novices and experts alike. Standard design conventions of CQAs, search engines, and traditional frequently asked questions (FAQ) services were carefully combined with well-known usability principles such as Nielson's heuristics and Fitt's law to minimize the gulf of execution and evaluation in the use of ACQUA [19], [20].

In particular, four user activities are supported by ACQUA: (a) asking a question, (b) scrolling through the list of similar questions clustered from the four CQAs, (c) viewing the predicted best answer to a selected question, and (d) viewing all answers to a question from the original CQA site. The home page of ACQUA is shown in Fig. 1.

Fig. 1. Home page of ACQUA.

The simple and unobtrusive home page of ACQUA prompts users to ask their questions. On entering a question, ACQUA computes the quadripartite graph-based clustering across the aggregated corpora of the four CQAs to return users the cluster of questions that are similar to the entered question. For example, on entering the question "*How much is the monthly income of a mortgage broker*", ACQUA redirects users to the question clustering results as shown in Fig. 2.

Fig. 2. Question clustering results in ACQUA.

Users can scroll through the list of similar questions obtained from the question clustering results (Fig. 2). The number of questions found in the quadripartite cluster is also indicated. Each clustered question is depicted through three components: question title, best answer snippet, and original CQA indicator. The question title acts as a link that can be clicked to retrieve the best answer predicted by ACQUA for the particular question. The best answer snippet, consisting of a single line snapshot of the predicted best answer, follows immediately below the question title. The original CQA indicator icon beside each question serves as a link to the source CQA site from which the question has been retrieved. The page showing the question clustering results is also accompanied by a drop-down option that allows users to filter the results based on their preferred CQAs.

For every question in the question clustering results that has received at least one answer, the corresponding predicted best answer can be viewed by clicking on the individual question. For example, clicking the last question

ISBN: 978-988-19251-8-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

"*How much money do mortgage makers make?*" in the question clustering results (Fig. 2) causes a pop-up window to open as shown in Fig. 3. However, if a question in the cluster had not received any answer in its source CQA, the question title would not be clickable, and the snippet would indicate "*No answer received yet*". The answer prediction functionality of ACQUA uses a voting system to seek users' feedback ('*thumbs up*' for endorsement, and '*thumbs down*' for disapproval) on whether the predicted best answer could effectively meet their information needs. Such votes are treated as social features for answer prediction, and are also used to re-compute the best answer result for subsequent accesses. However, ACQUA does not display the number of '*thumbs up*' and '*thumbs down*' votes received by the predicted best answers to prevent users from hopping on the bandwagon.

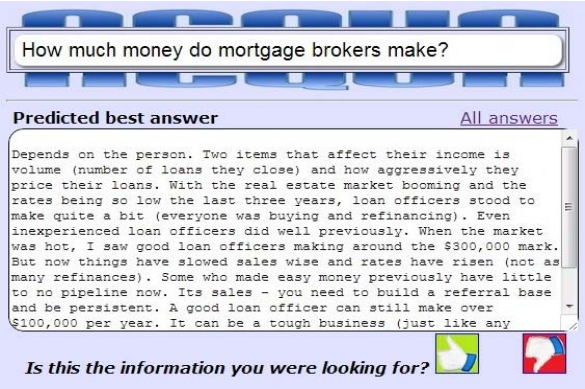

Fig. 3. Pop-up window showing answer prediction results in ACQUA.

In order to view all answers to a particular question from the original CQA site, users are provided with two options. For one, users can click on the original CQA indicator icon that appears beside every question in the question clustering results (Fig. 2). Additionally, they can click on the "*All Answers*" link that is provided in the pop-up window for answer prediction (Fig. 3). This feature of redirecting users to the original CQAs has been included in ACQUA given that some users may prefer to use actual CQAs in order to browse through all answers attracted by a given question. Moreover, they may also be interested to look through the comments received by such answers in the CQAs. Hence, ACQUA does not discourage users from accessing the original CQAs. It also does not take a toll on the integrity and uniqueness of the individual sites.

IV. Methodology

A. Data collection Methods

Data was collected using a combination of questionnaires and group interviews. The use of questionnaires allowed obtaining a broad overview of users' behavioral intention to adopt ACQUA [21]. Specifically, qualitative responses were sought through open-ended questions. Such an approach facilitates gathering rich and naturalistic data in a direct fashion, and is particularly suited to study users' information seeking behavior [22].

Semi-structured group interviews were conducted as a follow-up study to corroborate the findings from the questionnaires. Group interviews were chosen given their ability to elicit deeply nuanced views from participants. They are highly effective to explore participants' cognitive heuristics, and unearth comments that might not otherwise be raised through the sole use of questionnaires [23].

B. Participants

For the evaluation of ACQUA, a total of 42 participants (20 females) were recruited. A convenience sample was used by inviting volunteers to participate via e-mails. Of these, 24 were working professionals mostly hailing from the IT background. The rest were full-time graduate students from an institute of higher learning in Singapore. The average age of the participants was found to be 29 years (SD = 4.64, Min = 23, Max = 40).

Data collected from the participants indicated that all were frequent and active web surfers. Their average self-reported Internet experience was 4.52 out of 5 (SD = 0.75, Min = 4, Max = 5), with 1 being novice, and 5 being expert. All participants indicated uniform familiarity with the use of search engines and FAQ services. They were cognizant of the working of CQAs. In spite of not being registered CQA users, participants indicated familiarity with interfaces of CQAs such as Yahoo! Answers and WikiAnswers. This is conceivable as these CQAs are widely indexed by search engines in response to users' natural language queries [24].

C. Procedure

The evaluation of ACQUA was conducted in an institution of higher learning in Singapore over a period of two weeks in March, 2012. The 42 participants were randomly divided into six groups of seven members each. Every group of participants independently went through an evaluation session.

On an average, each evaluation session lasted for some 45 minutes, and was conducted in three steps. First, the participants were introduced to the prototype of ACQUA. To help them understand the features of ACQUA, the four usage scenarios were demonstrated: (a) asking a question, (b) scrolling through the list of similar questions clustered from the four CQAs, (c) viewing the predicted best answer to a selected question, and (d) viewing all answers to a question from the original CQA site.

Second, the participants were required to answer a questionnaire, which included questions pertaining to perceived usefulness, perceived ease of use, competence, and usability offered by ACQUA. In addition, some generic questions catering to the appeal of the prototype, as well as its drawbacks were also inquired. The response rate for the questionnaire was 100 %.

Third, the participants were organized in group interview sessions. Each group interview session lasted for some 15 minutes on an average. The interviewer asked questions pertaining to the perceived usefulness, perceived ease of use, competence, and usability of ACQUA. Findings from the focus groups were not impaired by any attrition

ISBN: 978-988-19251-8-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

problem.

V. Results

Most participants perceived ACQUA a useful prototype. The concept of aggregating questions and answers from multiple CQAs was unanimously appreciated. In particular, participant 42 was pleasantly surprised and commented "*OMG, So many CQAs at one location! That seems quite useful.*" Participants 3 and 39 felt that ACQUA could be quite useful for seeking information "*urgently*" and "*promptly*" given its ability to "*summarize information from multiple CQAs.*" Participant 12 indicated "*ACQUA saves a lot of time as you need not go to separate URLs for accessing different sites.*" The option to filter the question clustering results by specific CQAs was also lauded by all participants.

On perceived ease of use, majority of the participants agreed that ACQUA has a relatively simple interface, and is easy to use. Participant 15 appreciated "*the minimalist design*" of ACQUA. Participant 40 claimed that "*…the interface is so clean and uncluttered that there is perhaps no learning curve at all for an average Internet user*." Participant 41 complied by suggesting "*simplicity seems to be the USP of ACQUA.*" Given its perceived ease of use, most participants indicated that posting a question in ACQUA to retrieve answers from multiple CQAs can be much less cumbersome than going to the individual CQAs.

Most participants felt that use of ACQUA does not require any user competence. In particular, participant 28 stated "*using ACQUA is the same as using Google – both have equally easy and self-explanatory interfaces.*" Participant 13 indicated that "*ACQUA seems to have been designed for all users*," and its use does not call for "*any specialized training for an ordinary web user*". Furthermore, participants did not perceive the use of ACQUA monotonous. In fact, participants 2 and 19 claimed that using ACQUA will be "*…really engaging for all social media enthusiasts*" as there are "*…few sites that allow users compare results across multiple sites so efficiently.*"

In terms of usability, the consensus was not completely in favor of ACQUA. While around 70 % of the participants liked the usability of ACQUA due to its "*simple*" and "*clean*" interface, a few participants complained the lack of its flexibility as it does not provide users with ample scope for "*personalization*" and "*customization*". Moreover, participants 1, 29 and 35 wished ACQUA could present questions in the question clustering results in "*order of relevance*", as in search engines. Participant 20 further complained that ACQUA provides "*exit points to go out*" to other CQAs, but "*there are no entry points back to ACQUA – except tediously clicking the back button.*"

In terms of the general appeal, participants liked the concept of accessing multiple CQAs through a single interface. However, the most common dislikes for ACQUA can be attributed to the presentation of question clustering results. In its present form, ACQUA displays questions in alphabetical order of the source CQAs. The users disliked having no control on the ways questions could be floated to the top. Besides, few participants felt that ACQUA would neglect the integrity of the individual CQAs. In particular, participant 39 pointed that "*Most regular CQA users have definite preferences for sites.*" It might be difficult for such users to "*discard their preferred CQAs*" in favor of ACQUA.

VI. Discussion

Arising from the results, three findings could be culled. First, the behavioral intention to adopt ACQUA appears generally promising. Consistent with prior research [15], [25], [26], performance expectancy and effort expectancy emerged as the two crucial determinants of behavioral intention to adopt ACQUA. Performance expectancy is the extent to which users find an application useful to achieve a desired performance. On the other hand, effort expectancy refers to the degree to which users find an application engaging to use. While the former is pre-dominantly derived from perceived usefulness, the latter is significantly driven by perceived ease of and user competence [15], [27]. Fair responses received from majority of the participants on perceived usefulness, perceived ease of use, and user competence testifies that performance expectancy and effort expectancy are significant antecedents for behavioral intention to adopt.

Second, question clustering and answer prediction functionalities were appreciated by most participants. With respect to question clustering, participant 8 felt that retrieving answers would be easier as "*the same question may be answered in one CQA, and unanswered in the other*." This finding is in line with [28], [29], suggesting that question clustering can allow efficient reuse of CQA corpora in answering users' present questions with past answers. On the other hand, the main reason for the favorable response towards the answer prediction functionality was that most participants indicated reluctance to trust users' votes to select a good answer. Indeed, users' votes for answer prediction may not always be an accurate predictor of answer quality given such votes are largely voluntary, subjective, and hence, can be fallacious [1], [30].

Third, with web technologies becoming increasingly user-friendly, online users have become more sophisticated and tech-savvy in their information seeking behaviors [31]. Especially, they seem to have developed strong preferences for aggregation and customization. With increase in the number of websites, users are increasingly looking for aggregation and easy dissemination of information from multiple sites [32], [33]. With respect to customization, some participants complained the lack of flexibility in the order of the question clustering results. Besides, participant 33 indicated that "*…ACQUA cannot be customized to add a new question in the corpus; also no scope to answer a question.*" Ample scope for customization appears essential to engage customers in using an application [34]. Being spoilt for choices in terms of technology, modern users' behavioral intention to adopt seems to be strongly driven by quest for such sophistications.

ISBN: 978-988-19251-8-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

VII. Conclusion

This paper introduces ACQUA, a prototype for an aggregated CQA site. Aggregating the corpora of four CQAs, namely, Answerbag, Askville, WikiAnswers and Yahoo! Answers, ACQUA represents a system that incorporates both question clustering and answer prediction functionalities. Using a combination of questionnaires and group interviews, an evaluation of users' behavioral intention to adopt ACQUA is conducted based on four parameters, namely, perceived usefulness, perceived ease of use, competence, and usability. The results indicate that the behavioral intention to adopt ACQUA was generally promising. In particular, ACQUA received favorable responses from most participants in terms of perceived usefulness, perceived ease of use, and competence. In terms of usability however, the consensus was not completely in favor of ACQUA.

This paper needs to be viewed in light of two constraints. First, since ACQUA is a conceptual prototype, the participants could not play around with a fully functional version of the system. Second, due to the small cohort of participants, quantitative analysis was not conducted to avoid steep variations in responses. Nonetheless, the use of qualitative approach allowed for rich and naturalistic data to be obtained.

With the observed optimism in the participants' behavioral intention to adopt ACQUA, future work could consider development of the prototype into a fully functional version. For the real time crawling of multiple CQAs, different novel web traversal strategies such as board forum crawling [35] and intelligent crawling [36] might be experimented. Subsequently, suggestions provided during this evaluation study could be incorporated into the design.

References

[1] C. T. Lee, E. M. Rodrigues and G. Kazai, "Model for voter scoring and best answer selection in community Q&A services," in *International Joint Conferences on Web Intelligence and Intelligent Agent Technologies (WI-IAT '09)*, Milano, Italy, 2009, pp. 116-123.

[2] P. Shachaf and H. Rosenbaum, "Online social reference: A research agenda through a STIN framework," in *Proceedings of the iConference*, Chapel Hill, North Carolina, United States, 2009.

[3] D. Beeferman and A. Berger, "Agglomerative clustering of a search engine query log," in *Proceedings of the 6th ACM International Conference on Knowledge Discovery and Data Mining*, Massachusetts, United States, 2000, pp. 406-416.

[4] Y. Chen and F. A. Fonseca, "A bipartite graph co-clustering approach to ontology mapping," in *Workshop of Semantic Web Technologies for Searching and Retrieving Scientific Data*, Florida, United States, 2003.

[5] M. Rege, M. Dong and J. Hua, "Graph theoretical framework for simultaneously integrating visual and textual features for efficient web image clustering," in *Proceeding of the International Conference on World Wide Web*, Beijing, China, 2008, pp. 317-326.

[6] E. Agichtein, C. Castillo, D. Donato, A. Gionis and G. Mishne, "Finding high-quality content in social media," in *Proceedings of the International Conference on Web Search and Web Data Mining*, New York, United States, 2008, pp. 183-194.

[7] J. Bian, Y. Liu, E. Agichtein and H. Zha, "Finding the right facts in the crowd: Factoid question answering over social media," in *Proceedings of the International Conference on World Wide Web*, Beijing, China, 2008, pp. 51-60.

[8] P. Borlund, "The concept of relevance in IR," *Journal of the American Society for Information Science and Technology*, vol. 54, no. 10, pp. 913–925, May 2003.

[9] M. J. Blooma, A. Y. K. Chua and D. H. Goh, "Quadripartite graph-based clustering of questions," in *Proceedings of the 7th International Conference on Information Technology: New Generations*, Nevada, United States, IEEE Computer Society, 2011.

[10] M. J. Blooma, A. Y. K. Chua and D. H. L. Goh, "What makes a high-quality user-generated answer?," *IEEE Internet Computing*, vol. 15, no. 1, pp. 66-71, Jan.-Feb. 2011.

[11] R. Lambiotte and M. Ausloos, "Collaborative tagging as a tripartite network," Computational Science ICCS 2006, pp. 1114–1117.

[12] P. Katerattanakul and K. Siau, "Measuring information quality of web sites: Development of an instrument," in *Proceedings of the 20th International Conference on Information Systems*, North Carolina, United States, 1999, pp. 279– 285.

[13] Y. Guo and S. Barnes, "Why people buy virtual items in virtual worlds with real money," *The Data Base for Advances in Information Systems*, vol. 38, no. 4, pp. 70-76, Nov. 2007.

[14] C. Lopez-Nicolas, F. J. Molina-Castillo and H. Bouwman, "An assessment of advanced mobile services acceptance: Contributions from TAM and diffusion theory models," *Information & Management*, vol. 45, no. 6, pp. 359–364, Sep. 2008.

[15] A. Y. K. Chua, R. S. Balkunje and K. Chang, "Rendezvous - a social web-based application for knowledge sharing and entertainment," in *Proceedings of the International MultiConference of Engineers and Computer Scientists*, vol. 1, 2012.

[16] R. M. Peters, M. Janssen and T. M. van Engers, "Measuring e-government impact: existing practices and shortcomings," in *Proceedings of the 6th International Conference on Electronic Commerce*, Delft, Netherlands, 2004, pp. 480-489.

[17] F.D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Quarterly*, vol. 13, no.3, pp. 319-339, Sep. 1989.

[18] Y. Kim, S. Oh and H. Lee, "What makes people experience flow? Social characteristics of online games," *International Journal of Advanced Media & Communication*, vol. 1, no.1, pp. 76-92, Sep. 2005.

[19] J. Nielsen, *Usability Engineering*, San Francisco: Morgan Kaufmann, 1993, ch. 2.

[20] D. Norman, *The design of everyday things*, New York: Doubleday, 1988, ch. 2.

[21] K. Taylor & S. Corrall, "Personalized service? Changing the role of the government librarian," *Journal of Information Science*, vol. 33, no. 3, pp. 298-314, Jun. 2007.

[22] T.D. Wilson, "On user studies and information needs," *Journal of Documentation*, vol. 37, no. 1, pp. 3-15, 1981.

[23] J. M. Metzger, A. J. Flanagin and R. B. Medders, "Social and heuristic approaches to credibility evaluation online," *Journal of Communication*, vol. 60, no. 3, pp. 413–439, Sep. 2010.

[24] D. Roussinov, W. Fan and J. Robles-Flores, "Beyond keywords: Automated question answering on the web," *Communications of the ACM*, vol. 51, no. 9, pp. 60-65, Sep. 2008.

[25] O. Nov, C. Ye, "Resistance to change and the adoption of digital libraries: An integrative model," *Journal of the American Society for Information Science and Technology*, vol. 60, no. 8, pp. 1702–1708, Aug. 2009.

[26] C. M. Chiu and E. T. G. Wang, "Understanding web-based learning continuance intention: The role of subjective task value," *Information & Management*, vol. 45, no. 3, pp. 194-201, Apr. 2008.

[27] V. Venkatesh, M.G. Morris, G.B. Davis and F.D. Davis, "User acceptance of information technology: Toward a unified view," *MIS Quarterly*, vol. 27, no.3, pp. 425-478, Sep. 2003.

[28] A. Shtok, G. Dror, Y. Maarek and I. Szpektor, "Learning from the past: answering new questions with past answers," in *Proceedings of the International Conference on World Wide Web*, ACM, 2012, pp. 759-768.

[29] D. Carmel, M. Shtalhaim and A. Soffer, "eResponder: Electronic question responder," *Lecture Notes in Computer Science*, vol. 1901, pp. 150-161, Springer Berlin/Heidelberg, 2000.

[30] P. Fichman, "A comparative assessment of answer quality on four question answering sites," *Journal of Information Science*, vol. 37, no. 4, pp. 1-11, Oct. 2011.

[31] V. Kitzie and C. Shah, "Faster, better, or both? Looking at both sides of online question answering coin," in Proceedings of the American Society for Information Science and Technology, vol. 48, 2011, pp. 1-4.

[32] T. Gupta, S. Garg, N. Carlsson, A. Mahanti and M. Arlitt, "Characterization of Friendfeed: A web based social aggregation

service," in *Proceedings of the International Conference on Weblogs and Social Media*, San Jose, United States, 2009.

[33] B. Jingyuan and G. Xingwu, "Situation and development countermeasures of Chinese government website," in *Proceedings of the International Conference on Artificial Intelligence, Management Science and Electronic Commerce*, 2011, pp. 5777-5780.

[34] W. T. Boon and T. W. Lob, "The impact of interface customization on the effect of cognitive style on information system success," *Behaviour & Information Technology*, 10(4), pp. 297-310, 1991.

[35] Y. Guo, K. Li, K. Zhang and G. Zhang, "Board forum crawling: A web crawling method for web forum," in *Proceedings of the International Conference on Web Intelligence*, Washington, United States, 2006, pp. 745-748.

[36] R. Cai, J. M. Yang, W. Lai, Y. Wang and L. Zhang, "iRobot: An intelligent crawler for web forums," in *Proceedings of the International Conference on World Wide Web*, Beijing, China, 2008, pp. 447-456.

ISBN: 978-988-19251-8-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)